

## WHAT IS CLAIMED

1. An audio decoding device including a jitter buffer for storing a received packet, and decoding means for decoding the packet stored in the jitter buffer,  
the audio decoding device comprising:  
    playback speed change means for changing, with respect to a decoded audio signal obtained by the decoding means, the playback speed thereof;  
    an output buffer for temporarily storing a digital audio signal outputted from the playback speed change means;  
    means for reading out the digital audio signals stored in the output buffer at predetermined time intervals;  
    playback speed control means for controlling the playback speed change means on the basis of the number of packets stored in the jitter buffer; and  
    decoding timing control means for controlling the timing of decoding by the decoding means on the basis of the amount of data stored in the output buffer.
2. The audio decoding device according to claim 1, characterized in that the playback speed control means controls the playback speed change means such that the playback speed is reduced when the number of packets stored in the jitter buffer is less than a first predetermined reference value, while controlling the playback speed change means such that the playback speed is increased when a state where the number of packets stored in the jitter buffer is more than a second predetermined reference value which is not less than said first predetermined reference value is continued for a predetermined time period.
3. The audio decoding device according to either one of claims 1 and 2,

characterized in that

the decoding timing control means requires the decoding means to decode the packet when the amount of data stored in the output buffer is less than a predetermined reference value.

4. An audio decoding device including a jitter buffer for storing a received packet, and decoding means for decoding the packet stored in the jitter buffer,

the audio decoding device comprising:

delay time control means for carrying out such control that a delay time period elapsed from the time when the packet is stored in the jitter buffer until the packet is decoded is lengthened when the number of packets stored in the jitter buffer is less than a first predetermined reference value, while carrying out such control that a delay time period elapsed from the time when the packet is stored in the jitter buffer until the packet is decoded is shortened when a state where the number of packets stored in the jitter buffer is more than a second predetermined reference value which is not less than the first predetermined reference value is continued for a predetermined time period.

5. The audio decoding device according to claim 4, characterized in that the delay time control means comprises

playback speed change means for changing, with respect to a decoded audio signal obtained by the decoding means, the playback speed thereof,

an output buffer for temporarily storing a digital audio signal outputted from the playback speed change means,

means for reading out the digital audio signals stored in the output buffer at predetermined time intervals, and

means for controlling the playback speed change means such that the

playback speed is reduced when the number of packets stored in the jitter buffer is less than the first predetermined reference value, while controlling the playback speed change means such that the playback speed is increased when a state where the number of packets stored in the jitter buffer is more than the second predetermined reference value which is not less than said first predetermined reference value is continued for a predetermined time period.

6. The audio decoding device according to claim 4, characterized in that the delay time control means controls the packet to be read out of the jitter buffer and fed to the decoding means such that the packet read out of the jitter buffer at the timing of packet reading is repeatedly decoded at the timing of packet reading continued a plurality of number of times including the current time, and the read-out of the packet from the jitter buffer is inhibited during the decoding when the number of packets stored in the jitter buffer is less than the first predetermined reference value, while controlling the packet to be read out of the jitter buffer and fed to the decoding means such that the plurality of packets stored in the jitter buffer are read out at a time at the timing of packet reading, and one of the packets is decoded and the other packets are discarded when the state where the number of packets stored in the jitter buffer is more than the second predetermined reference value which is not less than the first predetermined reference value is continued for a predetermined time period.

7. A network telephone set including a jitter buffer for storing a received packet, and decoding means for decoding the packet stored in the jitter buffer, the network telephone set comprising:  
playback speed change means for changing, with respect to a decoded

audio signal obtained by the decoding means, the playback speed thereof;

an output buffer for temporarily storing a digital audio signal outputted from the playback speed change means;

means for reading out the digital audio signals stored in the output buffer at predetermined time intervals;

playback speed control means for controlling the playback speed change means on the basis of the number of packets stored in the jitter buffer; and decoding timing control means for controlling the timing of decoding by the decoding means on the basis of the amount of data stored in the output buffer.

8. The network telephone set according to claim 7, characterized in that the playback speed control means controls the playback speed change means such that the playback speed is reduced when the number of packets stored in the jitter buffer is less than a first predetermined reference value, while controlling the playback speed change means such that the playback speed is increased when a state where the number of packets stored in the jitter buffer is more than a second predetermined reference value which is not less than said first predetermined reference value is continued for a predetermined time period.

9. The network telephone set according to either one of claims 7 and 8, characterized in that

the decoding timing control means requires the decoding means to decode the packet when the amount of data stored in the output buffer is less than a predetermined reference value.

10. A network telephone set including a jitter buffer for storing a received

packet, and decoding means for decoding the packet stored in the jitter buffer,  
the network telephone set comprising:

delay time control means for carrying out such control that a delay time period elapsed from the time when the packet is stored in the jitter buffer until the packet is decoded is lengthened when the number of packets stored in the jitter buffer is less than a first predetermined reference value, while carrying out such control that a delay time period elapsed from the time when the packet is stored in the jitter buffer until the packet is decoded is shortened when a state where the number of packets stored in the jitter buffer is more than a second predetermined reference value which is not less than the first predetermined reference value is continued for a predetermined time period.

11. The network telephone set according to claim 10, characterized in that the delay time control means comprises  
playback speed change means for changing, with respect to a decoded audio signal obtained by the decoding means, the playback speed thereof,  
an output buffer for temporarily storing a digital audio signal outputted from the playback speed change means,  
means for reading out the digital audio signals stored in the output buffer at predetermined time intervals, and  
means for controlling the playback speed change means such that the playback speed is reduced when the number of packets stored in the jitter buffer is less than the first predetermined reference value, while controlling the playback speed change means such that the playback speed is increased when the state where the number of packets stored in the jitter buffer is more than the second predetermined reference value which is not less than said first predetermined reference value is

continued for a predetermined time period.

12. The network telephone set according to claim 10, characterized in that the delay time control means controls the packet to be read out of the jitter buffer and fed to decoding means such that the packet read out of the jitter buffer at the timing of packet reading is repeatedly decoded at the timing of packet reading continued a plurality of number of times including the current time, and the read-out of the packet from the jitter buffer is inhibited during the decoding when the number of packets stored in the jitter buffer is less than the first predetermined reference value, while controlling the packet to be read out of the jitter buffer and fed to the decoding means such that the plurality of packets stored in the jitter buffer are read out at a time at the timing of packet reading, and one of the packets is decoded and the other packets are discarded when the state where the number of packets stored in the jitter buffer is more than the second predetermined reference value which is not less than the first predetermined reference value is continued for a predetermined time period.